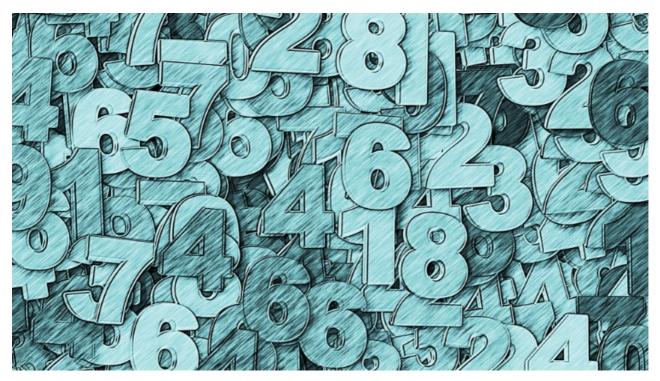# Descriptive statistics are essential to making complex analyses useful.

In response to the ever-growing volume of data, quantitative social research has become increasingly dependent on complex inferential methods. In this post, **Kevin R. Murphy** *argues that whilst these methods can provide insights, they should not detract from the significance of the comparatively simple descriptive statistics often found in table 1, which play an important role in communicating the significance of the research to policymakers and other research users.*

Research in the social and behavioural sciences increasingly relies on complex statistical methods to make sense of data. As researchers start to incorporate "big data" (i.e., very large data sets, often collected unobtrusively and automatically, characterized by the "four Vs" – i.e., volume, velocity, variety, veracity) into their research, the trend toward reliance on complex statistical methods to filter and analyse data is likely to increase. This increasing statistical sophistication has real benefits, allowing researchers to pursue questions that simpler and more traditional data analytic tools cannot fully address, but it also creates many challenges.

First, as statistical analyses become more complex, the likelihood that researchers or the consumers of research understand what these analyses can and cannot tell them decreases. Second, there is clear evidence that as statistical analyses become more complex, the likelihood of serious errors in analysis and interpretation increases. Third, many of the methods embraced by researchers in the social and behavioural sciences rely strongly and sometimes exclusively on a tool that is quickly losing support in the scientific community – i.e., null hypothesis significance testing. Key decisions in statistical analysis often rely on an assessment of whether particular parameters in a complex statistical model are statistically significant. Statistical significance tests are relatively easy to understand when applied to simple statistics (e.g., the correlation between two variables, the difference between a few means), but as analyses become more complex, the key components of statistical tests (e.g., the standard errors of the parameters of your model) become increasingly complex, making it increasingly difficult to understand *why* a particular parameter is significantly different from zero.

Perhaps the most important drawback of complex statistical methods is that they make it difficult for researchers to communicate simply and accurately with customers, clients, policy makers and others who use the results of our research to develop interventions or formulate policies. Instead of giving us better insights into human behaviour, complex statistical methods often do little more than confuse end users, bamboozle editors and reviewers, and create an impenetrable fog around our findings. Ironically, virtually every paper published in the behavioural and social sciences provides an essential tool for solving the problem of needless complexity in the analysis of data – i.e., Table 1, the table of descriptive statistics that is routinely presented and routinely ignored in virtually every published paper and report.

> Perhaps the most important drawback of complex statistical methods is that they make it difficult for researchers to communicate simply and accurately with customers, clients, policy makers and others

I am an Organizational Psychologist, and over the last 40 years I have reviewed I have reviewed several thousand papers and research reports. It is almost universal practice in my field for authors to mention Table 1 once (e.g., "Descriptive statistics are presented in Table 1) and to ignore it from that point forward. My interest in placing more emphasis on descriptive statistics can be traced in part to my experience reviewing several papers for top journals in which the ideas advanced by researchers and their interpretations of data were clearly impossible given the descriptive statistics shown in their Table 1. For example, I reviewed a well-reasoned paper dealing with the way organizations respond to crises created by resource shortages in which Table 1 made it clear that virtually none of the organizations included in the studies were ever short of resources. I have reviewed numerous papers claiming that some variable Z mediates the relationship between two other variables, X and Y, when Table 1 makes it clear: (a) X is not related to Y, meaning that there is nothing to mediate, or (b) Z had nothing to do with X or with Y, meaning that it cannot possibly work as a mediator. My experience as a researcher, a reviewer and a journal editor has led me to believe that the importance of tables in a research report is inversely related to the table number. That is, Table 1 is the most important table, not the one that should be mentioned and ignored, and the further into the weeds you get with complex analyses (e.g., Table 10), the less likely it is that whatever you have to say will both be correct and interpretable.

Descriptive statistics should serve two critical roles in every study we publish. First, they are a primary tool for communication. We live in a Golden Age of data visualization; the data analysis languages (e.g., **R**) and statistical packages we routinely use provide powerful but underutilized tools for visualizing data and for communicating what data mean. If you hope to communicate effectively with policy makers and with users outside of your academic field, good graphics are your best friend. More important, descriptive statistics should serve a gatekeeping role. That is, before you launch into a complex statistical analysis, you should always do what you can to determine whether your ideas are plausible, and whether they can plausibly be tested with the data at hand. In a [recent paper](#), I proposed that "Any result that is established based on a complex data analysis that cannot be shown to be at least plausible based on the types of simple statistics shown in Table 1 (e.g., means, standard deviations, correlations) should be treated as suspect and interpreted with the utmost caution." (p. 467). Complex statistical analyses can yield useful insights, but if you cannot show that your ideas are at least possible based on the information in Table 1, you face two problems. First, you might be wrong, perhaps seriously wrong in your interpretation of the data, and there may be no way to tell whether you are wrong or right. Second, you will find it difficult to communicate with non-expert audiences. The more serious attention you pay to Table 1, the better your research and the more likely it is that your research will have meaningful impact. It is time to give descriptive statistics their due!

This post draws on the author's paper, [In praise of Table 1: The importance of making better use of descriptive statistics](#), published in Industrial and Organizational Psychology: Perspectives on Science and Practice and book chapter, [Surviving the statistical arms race](#). In K.R. Murphy (Ed). Data., Methods and Theory in the Organizational Sciences: A New Synthesis. New York: Routledge.

Image Credit: [Geralt](#) via Pixabay.

Impact of Social Sciences Blog: Descriptive statistics are essential to making complex analyses useful.

Page 3 of 3